



遗传学 (第3版)

第19章 群体遗传与进化

1. 群体的遗传组成
2. Hardy-Weinberg定律
3. 影响群体遗传平衡的因素
4. 自然群体中的遗传变异及其检测
5. 物种及物种形成
6. 中性突变与分子进化
7. 新基因和蛋白质功能的起源
8. 人类进化概述





达尔文进化学说有3个重要原则：第一，**变异**的原则：在任何一个群体中的不同个体间都存在形态、生理和行为上的差异。第二，**遗传**的原则：后代与其亲本的相似性大于与其无关个体的相似性。第三，**选择**的原则：在特定的环境下，一些类型的个体总会比另一些类型的个体有更强的生存和繁殖能力。

群体遗传学（population genetics）是应用数学和统计学方法研究群体中的基因及其频率和可能的基因型及其频率以及影响这些频率的选择效应、突变作用、迁移和遗传漂变作用与群体遗传组成（genetic composition）的关系，探讨生物微（观）进化（microevolution）的机制。

群体遗传学是生物进化的重要的理论基础。





19.1 群体的遗传组成

19.1.1 孟德尔群体与基因库

(1) **群体** (population) :

享有一个共同的基因库，并能相互交配的一群个体。

(2) **孟德尔式群体** (Mendelian population) :

群体内的个体享有共同的基因库，能够相互交配，通过有性生殖传递基因，可用孟德尔定律进行分析的群体。一个最大的孟德尔群体就是一个物种。

(3) **基因库** (gene pools) :

有性生殖生物的一个群体中，能进行生殖的所有个体所携带的全部基因或遗传信息。



19.1.2 基因频率与基因型频率

群体中遗传着的基因及其频率以及可能的基因型及其频率构成了一个特定群体的遗传组成。研究群体的遗传结构变化的机制是群体遗传学的宗旨。

等位基因频率 (alleles frequency)：在一个二倍体的某特定基因座上某一个等位基因占该座位上等位基因总数的比率定义为该等位基因的频率。

$$\text{基因频率} = \frac{(2 \times \text{某一基因型纯合子个体总数} + \text{杂合子总数})}{\text{群体等位基因总数}}$$

基因型频率 (genotype frequency)：群体中某特定基因型个体的数目，占个体总数目的比率。全部基因型频率的总和等于1。

$$\text{基因型频率} = \text{某一基因型个体总数} / \text{群体总数}$$



- 设：在 N 个个体的群体中有一对等位基因 A 、 a 在常染色体上遗传， A 基因的频率为 p ， a 基因频率为 q 。其可能的基因型及其频率分别为： AA 基因型频率为 D ； Aa 基因型频率为 H ； aa 基因型频率为 R 。

$$\text{有： } p + q = 1; \quad D + H + R = 1$$

可以证明： $P_A = D + 1/2 H$ ； $q_a = R + 1/2 H$

例如，在某次抽样检测我国北方汉族人群的9274人的MN血型，说明等位基因 L^M 及 L^N 频率及其基因型 $L^M L^M$ ， $L^M L^N$ ， $L^N L^N$ 频率的计算方法（表19-1）。

表 19-1 MN 血型的群体遗传分析

表型	基因型	人数	基因型频率	等位基因数目		等位基因总数
				L^M	L^N	
M 型	$L^M L^M (D)$	2 235	0.241	4 470		
MN 型	$L^M L^N (H)$	4 460	0.481	4 460	4 460	
N 型	$L^N L^N (R)$	2 579	0.278		5 158	
总数		9 274	1	8 930	9 618	18 548

根据定义，依表19-1的资料分别计算基因型及基因频率：

$$L^M L^M \text{ 基因型频率 } D = \frac{2\,235}{9\,274} = 0.241$$

$$\text{基因 } L^M \text{ 频率 } p = \frac{8\,930}{18\,548} = 0.48$$

$$\text{基因 } L^N \text{ 频率 } q = \frac{9\,618}{18\,548} = 0.52$$

$$q = 1 - p = 1 - 0.48 = 0.52$$

或根据公式求得：

$$p_{L^M} = D + \frac{1}{2}H = 0.241 + \frac{1}{2} \times 0.481 = 0.241 + 0.2405 = 0.48$$

$$q_{L^N} = R + \frac{1}{2}H = 0.278 + \frac{1}{2} \times 0.481 = 0.278 + 0.2405 = 0.52$$

$$\text{且必有: } p + q = 0.48 + 0.52 = 1$$

$$D + H + R = 0.241 + 0.481 + 0.278 = 1$$





19.2 Hardy-Weinberg定律

19.2.1 Hardy-Weinberg定律的内容

“在一个大的随机交配的孟德尔群体内，**基因型频率**同时在**没有选择**（selection）、**没有突变**（mutation）、**没有迁移**（migration）和**遗传漂变**（genetic drift）发生的理想条件下，**世代相传保持不变**。”此定律是在1908年分别由英国数学家G.H.Hardy和德国医生W.Weinberg各自独立地发现的，故称哈迪-温伯格定律（law of Hardy-Weinberg），亦称哈迪-温伯格平衡（Hardy-Weinberg equilibrium）。



Hardy—Weinberg Law解释了繁殖如何影响群体的基因和基因型频率。此定律是群体遗传学的重要理论基石，也是进化理论的基础。

Hardy-Weinberg定律的要点是：

① 在随机交（婚）配下的孟德尔群体中，若没有其他因素（基因突变、选择、迁移和遗传漂变）的干扰，基因频率不变。

② 无论群体的起始状况如何，常染色体上的一对等位基因的基因型频率的平衡由下列二项式的展开式所决定： $(pA+qa)^2 = p^2(AA) + 2pq(Aa) + q^2(aa)$ ，达到平衡的速度只需一个世代的随机交配。

③ 如果随机交配系统得以保持（这是平衡定律最关键的假设），群体中的基因型频率保持在上述平衡状态不会改变，这便是理想群体中基因的遗传行为。



例：用两个等位基因说明Hardy-weinbery定律

		雌配子频率	
		P (A)	q(a)
雄配子频率	P(A)	$P^2(AA)$	$pq(Aa)$
	q(a)	$pq(Aa)$	$q^2(aa)$

由此可见这种随机婚配子代的基因型和它们的频率是：

$$p^2[AA]+2pq[Aa]+q^2[aa]$$

子I代群体所产生的A 基因配子和a 基因配子的频率是：

$$A=P^2+1/2(2pq)=p^2+pq=p(p+q)$$

$$a=q^2+1/2(2pq)=q^2+pq(p+q)$$





现在来分析子₁代群体婚配。子₁代群体和亲代群体一样有3种基因型。这3种基因型的个体在随机婚配的情况下互相婚配的机会也可由概率乘法定理来推导，见下图。

		男 性 群 体		
		AA p ²	Aa 2pq	aa q ²
女 性 群 体	AA p ²	AA×AA P ⁴	AA×Aa 2p ³ q	AA×aa p ² q ²
	Aa 2pq	Aa×AA 2p ³ q	Aa×Aa 4p ² q ²	Aa×aa 2Pq ³
	aa q ²	aa×AA p ² q ²	aa×Aa 2pq ³	aa ×aa q ⁴



男性群体与女性群体这3种基因型随机婚配共有6种不同的婚配类型，由上图可写出这6种婚配类型的频率，见下表。

6种婚配类型及其频率

婚配类型	频率	婚配类型	频率
AA x AA	p^4	AA x aa	$p^2q^2+p^2q^2=2p^2q^2$
AA x Aa	$2p^3q+2p^3q=4p^3q$	Aa x aa	$2pq^3+2pq^3=4pq^3$
Aa x Aa	$4p^2q^2$	aa X aa	q^4



表 7-2 子₁代婚配类型和子₂代基因型及其频率

子 ₁ 代婚配类型	婚配频率	子 ₂ 代基因型及频率		
		AA	Aa	aa
AA × AA	p^4	p^4	—	—
AA × Aa	$4p^3q$	$2p^3q$	$2p^3q$	—
Aa × Aa	$4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
AA × aa	$2p^2q^2$	—	$2p^2q^2$	—
Aa × aa	$4pq^3$	—	$2pq^3$	$2pq^3$
aa × aa	q^4	—	—	q^4

子₂代基因型频率为：

$$\begin{aligned}
 & p^2(p^2 + 2pq + q^2)[AA] + 2pq(p^2 + 2pq + q^2)[Aa] + q^2(p^2 + 2pq + q^2)[aa] \\
 = & p^2(p + q)^2[AA] + 2pq(p + q)^2[Aa] + q^2(p + q)^2[aa] \\
 = & p^2[AA] + 2pq[Aa] + q^2[aa]
 \end{aligned}$$

由此可见，在子₂代群体中各种基因型的频率与子₁代群体完全相同。我们可以连续推算后代的基因型频率，结果都一致。所以就这对基因而言，在群体中处于遗传平衡状态。群体中基因型的相对比例保持不变，都是 $p^2+2pq+q^2$ 。这就是 Hardy-Weinberg 平衡数学公式，即纯合基因型频率是基因频率的自乘，其系数是1；杂合基因型的频率是两个有关基因频率的乘积，其系数是2。



19.2.2 平衡群体的特征及其应用

已经达到了Hardy-Weinberg平衡的群体，有5个基本特征，了解平衡群体的性质，有助于运用于遗传咨询和群体遗传分析（详见e19-1）。



19.2.3 χ^2 检验抽样群体中的基因型频率的平衡

判断某抽样群体中的基因型频率是否符合Hardy-Weinberg平衡，需要运用 χ^2 检验方法。

【例】在我国某大城市随机抽样调查1788人的MN血型，其中397人M型 ($L^M L^M$)；861人MN型 ($L^M L^N$)；530人是N型 ($L^N L^N$)。请问该3种基因型及其频率的分布是否符合Hardy-Weinberg平衡定律？

对于这类问题，首先需要从实际所收集的数据中计算 L^M 基因频率 (p) 和 L^N 基因频率 (q) (表19-2)。

表 19-2 MN 血型的基因频率计算

基因型(表型)	人数	L^M	L^N	总计
$L^M L^M$ (M型)	397 人	794		
$L^M L^N$ (MN型)	861 人	861	861	
$L^N L^N$ (N型)	530 人		1 060	
总数	1 788(N)	1 655	1 921	3 576

L^M 基因频率

$$p = \frac{397 \times 2 + 861}{3576} = \frac{1655}{3576} = 0.4628$$

L^N 基因频率

$$q = \frac{861 + 530 \times 2}{3576} = \frac{1921}{3576} = 0.5372$$

将计算得到的 p, q 值代入下表, 作 χ^2 检验的相关计算(表 19-3):

表 19-3 MN 血型基因型频率 χ^2 检验的计算

	$L^M L^M$	$L^M L^N$	$L^N L^N$	总计
实计频数(O)	397	861	530	1788(人)
预计频数(E)	(Np^2) 382.96	$(N2pq)$ 889.05	(Nq^2) 515.98	(N) 1788(人)
$\frac{(O-E)^2}{E}$	0.51	0.88	0.38	1.77

计算期望数:

$$np^2 = 1788 \times (0.4628)^2 = 1788 \times 0.2142 = 382.96$$

$$n2pq = 1788 \times 2 \times 0.4628 \times 0.5372 = 889.05$$

$$nq^2 = 1788 \times (0.5372)^2 = 1788 \times 0.2886 = 515.99$$



计算得 $\chi^2=1.77$ ，根据自由度 $df=3-1-1=1$ ， $P = 0.05$

查 χ^2 值表： $df = 1$ ， $\chi^2_{(1)0.05} = 3.841$ ([表3-3](#))

$\because \chi^2 < \chi^2_{0.05}$ (**$1.77 < 3.841$**)

$\therefore P > 0.05$ ($0.22 > p > 0.10$) **差异不显著。**

表明3种基因型频率符合Hardy-Weinberg平衡定律。

在此 χ^2 检验过程中注意自由度的确定，因为在计算预期频数时，要应用基因频率 p ，而它是以样本的实得数据中估计得来的，因此 χ^2 的自由度又减去1。



19.2.4 Hardy-Weinberg 定律的扩展

(1) 复等位基因的遗传平衡

如果有3个等位基因A、 a' 、 a 在群体中遗传，其频率分别为 p 、 q 、 r ，且 $p+q+r=1$ 。在一个大的随机交配群体中3个基因的频率与6种基因型频率如果有下列关系时：

$$(A, a', a)^2 \rightarrow AA, Aa', Aa, a'a', aa', aa$$

$$(p+q+r)^2 = p^2 + 2pq + 2pr + q^2 + 2qr + r^2$$

则认为这3个复等位基因的6种基因型频率的Hardy-Weinberg平衡已经建立。

平衡状态下的基因频率可以由基因型频率按下列各式求得：

$$p = p^2 + \frac{1}{2}(2pq + 2pr)$$

$$q = q^2 + \frac{1}{2}(2pq + 2qr)$$

$$r = r^2 + \frac{1}{2}(2pr + 2qr)$$

上述性质可以下列形式表示：

$$\begin{array}{cccccc}
 AA & \underline{Aa} & \underline{Aa'} & \underline{a'a'} & \underline{a'a} & aa \\
 p^2 + & 2p(q+r) + & (q+r)^2
 \end{array}$$

人类的ABO血型是由复等位基 I^A ， I^B 及 i 决定的。设：基因 I^A 的频率为 p ， I^B 基因频率为 q ， i 基因频率为 r 。在一个选择、突变、迁移等因素不起作用的、随机婚配的群体中，ABO血型及其频率的分布如下表（表19-4）。

表 19-4 人类 ABO 血型的表型与基因型频率

基因型	$I^A I^A, I^A i$	$I^B I^B, I^B i$	$I^A I^B$	ii
表型(频率)	A(\bar{A})	B(\bar{B})	AB(\bar{AB})	O(\bar{O})
预计的表型频率	$p^2 + 2pr$	$q^2 + 2qr$	$2pq$	r^2



平衡群体鉴定

•	AA	Aa	aa	是否平衡
1.	0.5	/	0.5	-
2.	0.4	0.2	0.4	-
3.	0.25	0.5	0.25	+

例3: $AA=0.25=P^2$ $p=0.5$

$aa = 0.25=q^2$ $q=0.5$

$(p+q)^2 = P^2+q^2+2pq=0.5^2+ 0.5^2+2 \times 0.5 \times 0.5 = 0.25+0.25+0.5$



如果一对等位基因有显性和隐性之分，杂合子的表现型与显性纯合子相同，这时基因频率和杂合子频率的估计值可利用Castle-Hardy-Weinberg平衡的数学公式得出。

设某一位点有等位基因D和d，纯合隐性基因型dd导致一种隐性遗传病。现在要问人群中一个个体得该种隐性遗传病的机会是多少。设D和d的基因频率分别为p和q。在实际工作中，隐性遗传病发病率 χ 可以通过群体调查知道，然后根据发病率求该隐性基因的频率。

隐性基因频率是发病率即隐性纯合子频率的平方根

$$q^2 = x \quad q = \sqrt{x}$$

由于 $p+q=1$ ， $p=1-q$ ，因此**显性基因频率是**： $p = 1 - \sqrt{x}$

我们还可以应用 Hardy-Weinberg 定律来估计杂合子的频率。

例：尿黑酸尿为常染色体隐性遗传病，约1 000 000儿童中有1个患儿，其发病率 $x=0.000001$ 。

按 Hardy-Weinberg 定律：

尿黑酸尿隐性基因频率： $q = \sqrt{x} = \sqrt{0.000001} = 0.001$

正常显性基因频率： $p = 1 - q = 1 - \sqrt{0.000001} = 0.999 \approx 1$

杂合子频率： $2pq = 2 \times 1 \times 0.001 = 0.002$

通过计算可见，当隐性基因频率很低时，则 p 接近于 1， $2pq$ 近似于 $2q$ ，此时杂合子的频率是隐性基因频率的 2 倍，即 $2q$ 。从计算可见人群中杂合子即携带者的频率要比患者频率高得多。杂合子频率与隐性纯合子频率之比为：

$$\frac{2pq}{q^2} = \frac{2p}{q} \approx \frac{2}{q}$$

即杂合子频率是隐性纯合子患者频率的 $2/q$ 倍。这意味着在隐性基因频率越来越小时，杂合子频率对隐性纯合子频率的倍数越高。



例：白化病人具aa基因型，正常人为AA或Aa，假定在某一群体中白化病人的频率为1/10000，按照Hardy-weinbery law:

纯合体aa的频率为 q^2

$$q^2=0.0001$$

$$q = \sqrt{0.0001} = 0.01$$

正常基因的频率: $1-0.01=0.99=P$

因此,基因型AA的频率: $p^2=0.99^2=0.98$

Aa为 $2pq=2 \times 0.99 \times 0.01 \approx 0.02$

(2) 伴性基因的遗传平衡

计算X-连锁座位上的基因频率较为复杂。在雌性纯合体中有两个相同的x-连锁等位基因，而在雌性异合体中也只有一个特定的等位基因，在所有雄性中只有一个X-连锁等位基因。为了在x-连锁座位上确定等位基因的数目，我们就得将雌性纯合体的数目乘2加上雌性杂合体数目和雄性半合体数目然后除以群体中等位基因的总数。确定等位基因总数时要用2倍的雌体数(因每个雌体都两个X-连锁基因座位)加上一倍的雄体数(因雄体只有单个X-连锁基因座位)

$$p = f(X^A) = \frac{(2X^A X^A \text{♀}) + (X^A X^a \text{♀}) + (X^A Y \text{♂})}{2 \times \text{雌体数} + \text{雄体数}}$$

例如有两个 X - 连锁的等位基因(X^A 和 X)可用下列等式来计算:

$$q = f(X^a) = \frac{(2X^a X^a \text{♀}) + (X^A X^a \text{♀}) + (X^a Y \text{♂})}{2 \times \text{雌体数} + \text{雄体数}}$$

X - 连锁座位上的基因频率也可由基因型频率来确定

$$p = f(X^A) = f(X^A X^A) + \frac{1}{2} f(X^A X^a) + f(X^A Y)$$

$$q = f(X^a) = f(X^a X^a) + \frac{1}{2} f(X^A X^a) + f(X^a Y)$$



例如：一对等位基因A、a在X染色体上遗传，在随机交配条件下，下列情况被认为达到了Hardy-Weinberg平衡，以通常的符号表示如下：

$$\left[\begin{array}{cc} \text{雄体} \\ (X^A & X^a) \\ p & q \end{array} \right] \text{ 和 } \left[\begin{array}{ccc} \text{雌体} \\ (X^A X^A & X^A X^a & X^a X^a) \\ p^2 & 2pq & q^2 \end{array} \right]$$

显然，伴X的基因处于平衡状态时，必定具有以下两个特征：

① 在雄性群体和雌性群体中的基因频率相等：

即 $p_X = p_{XX} = p$ ， $q_X = q_{XX} = q$ （下标X为雄性，XX为雌性）；

② 在雌性群体中，3种基因型频率必定有 $p^2 + 2pq + q^2$ 。可见，达到Hardy-Weinberg平衡时，雌体中的基因型频率与非伴性基因的基因型频率的平衡一样。而雄体中的基因型频率与其基因频率相同。



平衡群体中伴性基因频率的计算要从配子异型的性别入手。

人类的男性伴X基因的频率即为整个群体中该基因的频率。例如，计算随机婚配人群中红绿色盲基因的频率，即可根据男性中红绿色盲患者的调查数据来求。若发现色盲患者为7%，即表示 $q_x = q_{xx} = 0.07$ 。则可预期有 $q_{xx}^2 = 0.07^2 = 0.49\%$ 的女性患色盲。这一数字恰与女性群体中实际调查的发病率0.5%很接近。

运用平衡定律的理论可以作下列预见：对于隐性伴性性状而言，男性发病率：女性发病率 = $q : q^2$ 。对于伴X显性基因而言，则有男性发病率：女性发病率 = $p : (p^2 + 2pq) = 1 : (1 + q)$ ，可见女性发病率略高于男性。



19.3 影响群体遗传平衡的因素

19.3.1 非随机交配与莱特（Wright）定律

(1) 非随机交配

交配的随机性是Hardy-Weinberg定律的重要前提，若交配偏离，非随机的，则Hardy-Weinberg平衡关系就不复存在。

非随机交（婚）配的方式有两种：

①**近亲交配**或近亲结婚：有亲缘关系的个体间的交（婚）配。如在动物中常见的同胞兄妹、堂兄妹或人类中的表兄妹间的婚配。

②**选型交（婚）配**：是以交配个体的表型相似性为根据的交（婚）配方式。例如在人类中，选择与自己的身材或发色、肤色相似甚至智力水平相当的个体作为配偶。表型的选型婚配又可分为**正选型交（婚）配**和**负选型交（婚）配**又称选异交配。选择与自身表型特征不同的个体结合。

群体遗传学分析发现，**近亲交（婚）配**与**选型婚配**有相同的遗传效应，与随机交配下的基因型频率相比较时，前者将降低群体中杂合体频率，增加纯合体的频率。



(2) 赖特定律

赖特(S. Wright)于1922年提出了关于近亲繁殖群体的遗传平衡定律：当群体的基因频率(p, q)和近交系数(F)保持不变时，群体的基因型频率处于 $[p^2 + Fpq, 2pq(1 - F), q^2 + Fpq]$ 平衡状态。

赖特在研究近亲交（婚）配的群体中，用近交系数（coefficient of inbreeding, F ）来量化个体间近亲交（婚）配的程度，在人类群体中称近婚系数，其取值范围是从0~1， $F=0$ 表示没有近交， $F=1$ 表示完全的近交。

假设：一对等位基因 A, a ，各自的频率分别为 p, q 不变时，若群体中有某种程度的近交，近交系数为 F ，则 A 与 A 配子的结合频率不只为 p^2 ，它将大于 p^2 ， a 与 a 的结合频率也将会大于 q^2 ；如果每代的 F 不变，经证明群体中的基因型频率 D, H, R 保持在如下的平衡状态之中：

	D	H	R
	AA	Aa	aa
	$p^2 + Fpq$	$2pq - 2Fpq$	$q^2 + Fpq$
改写为:	$p^2 + Fpq$	$2pq(1 - F)$	$q^2 + Fpq$

该定律被誉为群体遗传学中另一重要的理论基石。该定律阐明了带近亲繁殖群体的遗传平衡原理；由于近交系数 F 的作用，若以 $0 < F < 1$ 正小数的 F 值为定值，每代作用于一对等位基因（ A, a ），处于Hardy-Weinberg平衡群体时，使群体中的 AA 基因型频率由 p^2 增加了 Fpq ， Aa 基因型频率由 $2pq$ 减少了 $2Fpq$ ， aa 基因型频率由 q^2 增加了 Fpq 。



对Wright定律几种情况的讨论：

① 当 $F=1$ 时，个体间进行高度的近交（或完全的近交），它是指发生在高等动植物群体中的自交（或自花授粉）。Wright定律有：

D	H	R
$p^2 + 1pq$	$2pq(1 - 1)$	$q^2 + 1pq$
p	0	q

一旦在一个Hardy-Weinberg平衡群体中实行自交，群体中的杂合体最终消失，将由一对等位基因 A 及 a 的基因频率（ p 与 q ）所代替。



② 当 $F=0$ 时，没有近交，即随机交配时，Wright平衡定律成为：

D	H	R
$p^2 + 0 \times pq$	$2pq(1 - 0)$	$q^2 + 0 \times pq$
p^2	$2pq$	q^2

可见，Hardy-Weinberg定律是当Wright定律中的 $F=0$ 时的特例。而Wright定律则可视作Hardy-Weinberg平衡的普遍化，无疑Wright定律阐明了基因在非随机交配群体中的一种遗传行为。由第6章已知，近交或远交只影响群体中的基因型频率，而不影响其基因频率。如果对该基因的某种纯合体（或杂合体）有不同的选择压时，近交就会间接地影响该群体后代中的基因频率。



(3) 近亲婚配的有害效应

群体遗传学分析说明，罕见的隐性遗传病在表（堂）亲婚配的家庭中的发生率比在随机婚配人群中的高出32.18倍。近亲婚配对人类的有害效应详见e19-3。

近亲婚配的危害主要表现为隐性遗传病纯合子患者的频率增高

表 19.1 表婚配和随机婚配出生隐性纯合子的频率

基因频率 (q)	随机婚配——隐性纯合子频率 (q ²)	表亲婚配——隐性纯合子频率 (q ² +pq/16)	Pq/16	两者之比
0.2	0.04	0.05	0.01	1.25
0.10	0.01	0.15625	0.005625	1.56
0.04	0.0016	0.004	0.0024	2.5
0.02	0.0004	0.001625	0.001225	4.06
0.01	0.00001	0.000719	0.000619	7.19
0.001	0.000001	0.0000635	0.0000625	63.5



19.3.2 基因突变与选择

在自然群体中突变、选择、迁移与漂变等因素影响群体遗传组成和调节遗传变异量，分析其作用是进一步理解群体遗传学的重要内容。

(1) 基因突变 (突变对基因频率的影响)

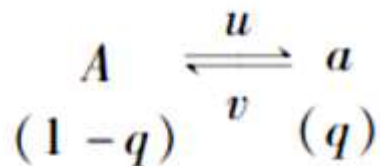
基因突变为进化提供了材料，是生物体进化机制的重要组成。

进化有两个阶段：一是遗传变异的产生；第二是不同等位基因频率的增加或减少来对进化压力作出反应。基因突变提供了遗传变异，这正是进化压力所作用的对象。

但是，基因突变频率很低，如果遗传的改变仅仅是突变，那么进化的速度将低到几乎不可能发生的程度。



如果群体中A与a基因的频率分别为p、q；由A→a的突变（正向突变）率为u，由a→A的回复突变（反向突变）率为v，则有如下反应：



由以上等式可见，由A→a的正突变会遭到a→A的对抗。每一代中有(1-q)u的A基因突变为a，有qv的a基因回复突变为A。若(1-q)u > qv，则a基因的频率增加，若(1-q)u < qv，则A基因的频率增加，于是每代a基因频率的净改变量为：

$$\Delta q = (1-q)u - qv$$

经过足够多的世代，这两种相反的力量相互抵消时， $\Delta q = 0$ ，群体处于平衡状态。在平衡时，a基因的平衡频率q可由下式求得：

$$(1-q)u = qv \quad u - qu = qv \quad \therefore q = \frac{u}{u+v}$$



- \hat{q} 与 \hat{p} 分别表示a与A基因平衡点的频率，平衡点的基因频率只决定于正、反突变的突变率，而与初始的基因频率无关，这是一种稳定的平衡。由于 $\hat{p}+\hat{q}=1$ ，故有：

$$\hat{q} = \frac{u}{u+v} \quad \hat{p} = \frac{v}{u+v}$$

设一个群体开始时的突变频率 $p = 0.9$ ， $q = 0.1$ 正向突变和回复突变的频率分别为：

$$u = 5 \times 10^{-5} \quad v = 2 \times 10^{-5}$$

则第一代基因频率的变化为：

$$\Delta p = vq - up = (2 \times 10^{-5} \times 0.1) - (5 \times 10^{-5} \times 0.9) = -0.000430$$

$f(A)$ 的减少仅0.043%。由于突变率非常低，因突变引起的基因频率的改变也非常低的



例：设某一真核生物 $A \rightarrow a$ 的突变率 $u = 10^{-5}$ ，而无回复突变的发生，则A的平衡频率是0。

若： $f(A)$ 从1 变为 0.99 需要1000代

从0.5 变为 0.49 需要2000代

从0.1 变为 0.09 需要10 000代

结论：通过突变使基因频率发生真正的改变需要很长的时间



(2) 自然选择 （自然选择对基因频率的影响）

自然选择(natural selection) 是在1858年由 C. Darwin 和 A. R. Wallace 在《进化论》中首先提出的。他们认为在自然界中，物种个体之间繁殖能力的不同导致个体具有不同的适应性。在任何生物群体中，最具竞争者具有最多的繁殖机会，即具有最多机会将其基因传给后代。

自然选择的作用：是指不同的遗传变异体具有差别的生活力和（或）差别的生（繁）殖力。

自然选择的本质就是一个群体中的不同基因型的个体对后代基因库作出的不同贡献。

自然选择是生物进化的动力。





① 适应值 (adaptive value)

度量自然选择作用的参数是达尔文适合度 (Darwinian fitness) 或适应值，是指在一定环境下，一种生物能够生存并把它基因传递到后代基因库中的能力。

适应值又称为相对适合度 (relative fitness)，一般用相对生育率 (fertility) 来表示：将具有最高生殖效能的基因型的适应值定为1，其他基因型与之相比较时的相对值，一般记作 w 。

例如：在人类群体中调查软骨发育不全侏儒症患者108人，其生育了27个子女，估计这些患者的相对生育率，则以他们的457个正常同胞总共生育了582个子女的生育率为对照，计算得到侏儒症的相对生育率即适应值为：

$$w = \frac{27/108}{582/457} = 0.2$$

这意味着每个正常人平均能留下1个子女，而侏儒症患者只能平均留下0.2个后代。



② 选择系数

与选择有关的另一个参数为**选择系数** (selective coefficient, S)，**即失去的生殖适合度**。选择系数是测量某一基因型在群体中不利于生存的程度，亦即在选择的作用下，被降低了的适合度，是自然选择的强度，即选择压 (selection pressure) 的度量， $S=1-w$ 。

显然适合度 $w = 1-S$

例如软骨发育不全患者的生育率 $w = 0.2$

则 $S = 1 - 0.2 = 0.8$

可见，负选择降低了生殖适合度

③ 选择对隐性纯合体的作用

对隐性体的不完全选择 即当 $0 < S < 1$ 的情况。

设：某基因座上有一对等位基因 A 、 a ，且 A 对 a 完全显性。3种可能基因型及其处于Hardy-Weinberg 平衡状态下的频率分别为： p^2 (AA) + $2pq$ (Aa) + q^2 (aa) = 1。设 AA 与 Aa 的适应值 $w = 1$ ，而隐性体 aa 的适应值为 $1 - S$ 。在选择前与选择后3种基因型的频率如表20 -8 所示。

表 20-8 显性完全,选择作用于隐性体时基因频率 q 的改变

基因型	AA	Aa	aa	全群体
起始频率	p^2	$2pq$	q^2	1
适应值 w	1	1	$1 - S$	\bar{w}
选择后	p^2	$2pq$	$q^2(1 - S)$	$1 - Sq^2$
相对频率	$p^2 / (1 - Sq^2)$	$2pq / (1 - Sq^2)$	$q^2(1 - S) / (1 - Sq^2)$	1



从上表可以看出**选择只对aa起作用**。

此时，隐性有害基因甚至是隐性的致死基因，即使没有新产生的突变来补偿，在选择压力下，仍然可以保留在许多世代中。

相对而言，选择对罕见的隐性个体的作用更无效。因为大多数隐性等位基因存在于杂合状态，自然选择对其不作用。平衡定律已指明，隐性基因频率越低，群体中杂合体的频率相对则越高。



对隐性体的完全选择 当 $S = 1$ 的情况：完全淘汰隐性体 aa ，令其在生殖年龄以前死去或是不育。这实质上是针对隐性致死基因的选择模型，是降低群体中隐性基因频率的最快速的一种方法。表20-9所示：

表 20-9 隐性体的完全淘汰

	基因型			总 和
	AA	Aa	aa	
选择前	p^2	$2pq$	q^2	1
适应值	1	1	0	
选择后	p^2	$2pq$	0	$p^2 + 2pq$
相对频率	$p^2 / (p^2 + 2pq)$	$2pq / (p^2 + 2pq)$	0	1



负选择对显性基因的作用最快，而对隐性基因的作用是比较缓慢的，因隐性致病基因多数存在于表现型正常的杂合子中。另外，在某些隐性遗传病中，在特定的条件下杂合子可能比正常纯合子个体更有利于生存而繁殖下一代，这称为**杂合子优势**(heterozygote advantage, heterosis)。例如在恶性疟疾地区镰状红细胞贫血、珠蛋白生成障碍性贫血和G6PD缺乏症的杂合子均有抗疟疾能力，从而在自然选择中获得生存的机会。

第五节 历史上的优生运动不符合群体遗传学原理

自从达尔文提出进化论以后,有些人的生物决定论的倾向很厉害,他们持这样一种思想:人类天生就有优有劣,这是环境无法改变的。从20世纪初到20世纪40年代,在遗传学发展之初,很多北美和欧洲国家就非常热衷于开展优生运动,以达到改良人种的目的。它们制定法律,那些被认为智力低下的人、犯罪的人要绝育。这一运动的顶点就是希特勒的“优生论”(eugenics),首先被拿来开刀的是本民族的残疾人,第一步是绝育,第二步是安乐死,因为注射效率太低,发明了毒气。随后,希特勒的屠刀又指向了所谓的“劣等民族”,如吉普赛人、犹太人。事实上是:一个社会不要梦想消除残疾人和那些患有严重的遗传病病人来达到“改良人种”的目的。即使现存的人都是健康的,还有基因突变的问题。何况有缺陷的人就一定是“劣生”的吗?梵高患有严重的精神病,坐在轮椅上的霍金是继爱因斯坦之后最伟大的物理学家,他们是“优生”还是“劣生”?撇开这些不谈,从 Castle-Hardy-Weinberg 平衡的角度而言,优生学也是完全站不住脚的。在优生学中一个最简单的问题,就是要消除由基因型 aa 引起的隐性遗传病。假定在一个随机婚配的大群体中,在某一位点上仅有一对等位基因 A、a, A 为正常显性基因, a 为隐性致病基因,它们的基因频率分别为 p、q。根据 Castle-Hardy-Weinberg 平衡推导,此群体的基因型频率分别为: p^2 、 $2pq$ 、 q^2 。根据优生学的法则,看看有害基因的频率 q 是如何随着致病的个体完全被消除而逐代下降的。在一个随机婚配的大群体中完全消除 aa 个体的选择模式由表 7-12 表示。

表 7-12 随机婚配群体中去除 aa 后基因频率的变化

基因型	频率	去除 aa	新的基因数	新基因频率
AA	p^2	p^2	} (A) $p^2 + pq = p$	$p' = 1/(1+q)$
Aa	$2pq$	$2pq$		
aa	q^2	0	} (a) $0 + pq = pq$	$q' = q/(1+q)$
合计	1.00	$p(p+2q) = p(1+q)$	$p + pq = p(1+q)$	1.00

上表共有五列,前两列仅显示了 Castle-Hardy-Weinberg 平衡的三种基因型的频率。第三列则显示了当所有的 aa 个体从群体中被消除后的群体基因构成。表面上看,在第三列中似乎因为 aa 个体消除后群体的人数减少,实质上取的是一个大的群体,aa 消除后余下的人数还是相当多的,通过频率的“标准化”,该群体又将成为一个新的整体。第四列则显示了 aa 个体消除后群体中的基因计数,所有的 AA 个体带的基因数,加上 Aa 个体带的基因数的一半,则构成了 A 基因数,而 aa 个体带的基因数,加上 Aa 个体带的基因数的一半,则构成了 a 基因的频率。表的最后一列则是把第四列的每一个基因数,除于经标准化后的总的基因数,从而得到新的基因频率。新的隐性基因频率用 q' 表示:

$$q' = \frac{pq}{p}(1+q) = \frac{q}{(1+q)} \quad (\text{公式1})$$

q 代表选择前的 a 基因频率, 而 q' 则表示经过一代选择后的 a 基因频率。若 $q = 1/2$, 则 $q' = (1/2)/(1+1/2) = 1/3$; $q = 1/20$, 则 $q' = (1/20)/(1+1/20) = 1/21$; $q = 1/200$, 则 $q' = (1/200)/(1+1/200) = 1/201$, 等等。

新 q' 的分母总是大于旧 q 的分母(在原来的分母上加 1), 所以 q' 的值, 随着选择代数的增加而逐代下降。假如想知道经十代选择后的结果, 只要将公式 1 重复运用十次即可, 这样很容易得到一个经任何代选择的通式, q 和 q' 作为世代数的标记, 用 q_0 和 q_1, q_2, \dots 分别表示原代和经一代、二代选择后的基因频率。 $q_1 = q_0/(1+q_0)$; $q_2 = q_1/(1+q_1)$ 等等, 则:

$$q_2 = \frac{q_0/(1+q_0)}{1+q_0/(1+q_0)} = \frac{q_0}{1+2q_0}$$

这样,对于经 n 代的选择来说,就可以得到:

$$\underline{q_n = q_0/(1+nq_0)} \quad (\text{公式 2})$$

那么需要多少代的选择才能使基因频率降到原先的 $1/2$? 即要知道 n 的值。设 $q_n = 1/2q_0$ 据公式 2, $q_0/(1+nq_0) = 1/2q_0$, 则 $nq_0 = 1$, $n = 1/q_0$ 。更进一步说,想知道需要经过多少代的选择才能使 q_0 降到 q_n 简单的办法就是利用公式 2, 从 q_0 和 q_n 来求出 n 的值:

$$q_n = (1+nq_0) = q_0$$

$$nq_0q_n = q_0 - q_n$$

$$\underline{n = 1/q_n - 1/q_0} \quad (\text{公式 3})$$

这个公式应用得更广泛,而且与公式 $n = 1/q_0$ 是完全一致的。在前面提到的特例 $q_n = 1/2q_0$ 中, $1/q_n = 2/q_0$, 所以 $n = 1/q_n - 1/q_0 = 2/q_0 - 1/q_0 = 1/q_0$, 这样就与公式 3 完全一致了。现在,对从每一代中完全消除 aa 个体后的 q 值作一估计。假定 $q_0 = 1/2$, 则经每一代选择后的 q 值为:

$$q_i = 1/2, 1/3, 1/4, 1/5, 1/6, 1/7, 1/8, 1/9, 1/10, \dots, 1/200, 1/201, 1/202, \dots, 1/400, 1/401, 1/402, \dots$$

从中可以看到 q 从 $1/2$ 降到 $1/4$ (是原值的一半) 需要两代的时间 ($1/2$ 的倒数), 从 $1/4$ 降到 $1/8$ 需要四代, 而从 $1/200$ 降到 $1/400$ 则需 200 代 ($1/200$ 的倒数)。200 代对人类历史来说意味着什么? 它几乎是耶稣基督公元开始至今的时间的 3 倍。可以用这样一句话来描述基因频率和频率下降所需时间的关系: 基因频率越低, 则频率下降所需的时间越长。 因为它们之间存在倒数关系 ($n = 1/q_0$)。 这还仅仅是数学上的推理, 实际情况远不是这样。当基因频率较高时, 如 $q = 1/2, 1/3, 1/4, \dots$ 时, 可以忽略突变, 因为此时突变的量值远远小于基因频率。然而, 对于一些较小的基因频率, 如 $q = 0.0075$ 时, 突变的量值就绝对不能忽略了。下面来看看突变是如何影响基因频率的: 假定正常基因 A 在每一代中有一部分将突变成隐性致病基因 a , 写成 $A \rightarrow a$, 显然, 这种突变将增加 q 的值, 尽管这种增加是很微小的。至少有两种因素能影响 q 的值: ① 选择将降低 q 的值; ② 突变将增加 q 的值。 这两种相反的力量最终将渐渐达到平衡状态。此时选择引起的 q 值降低, 与突变引起的 q 值增加相等, 这样 q 值将保持不变, 我们将找到这一平衡的 q 值, 用“ \hat{q} ”表示。它必然是突变速率与选择强度作用的结果。这个平衡值 (\hat{q}) 是稳定的。假定 $\hat{q} = 0.0035$, 若群体中 q 的实际值大于 0.0035 , 它将在下一代中下降; 相反, 若群体中 q 的实际值小于 0.0035 , 它将在下一代中增加。图 7-7

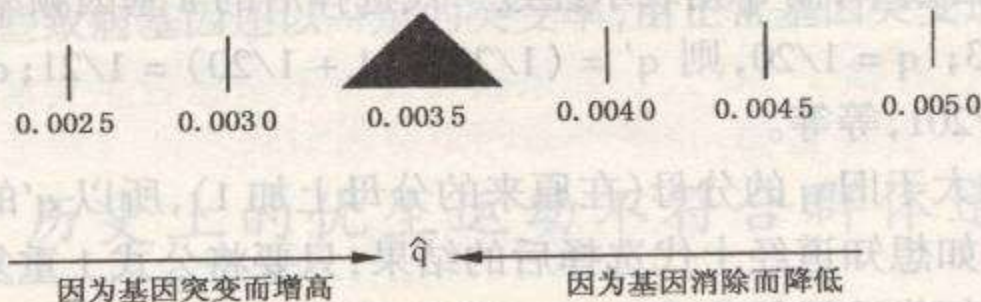


图 7-7 两种相反的因素对维持基因平衡稳定的影响

将帮助我们理解稳定平衡的基本概念。

我们看到以平衡点为中心(图中 $\hat{q} = 0.0035$), q 值在它左右摆动。让我们回到最初的例子,那就是从数学的推理来看,在群体中完全消除 aa 个体后,使基因频率从 0.0050 下降到 0.0025 需要 200 代的时间,现在可以清楚地看到,这是不可能的,在突变的作用下,它将在 $\hat{q} = 0.0035$ 这一点上停止不前,人们永远不能使 q 值达到 0.0025 ,更不用说使 q 值为 0 了。这就是历史上优生运动屡屡失败的原因,因为它不符合群体遗传学的科学原理。至于中国在人口问题上宣传的“优生优育”(well bear, well rear)是指生育健康,通过预防有遗传缺陷和先天性病患儿的出生,以达到提高人口遗传素质的目的。

(陆振虞)



(3) 突变与选择的联合效应

基因突变与自然选择对平衡群体基因频率的单独效应如果没有其他力量加以制止，将会在一个充分长的时间内逐渐将群体中的某种基因的频率由零改变成1，或是相反。然而在自然界中突变压和选择压如朝着同一方向同时起作用，改变基因频率的速度会比单独作用时更快。如果它们的作用彼此相反，效应可能相互抵消，因而对群体可能最终造成一种稳定的平衡（详见e19-5）。





19.3.3 迁移与遗传漂变

迁移 (migration) 或基因流 (gene flow) 是指个体从一个群体迁入另一群体或从一个群体迁出, 然后参与交配繁殖, 导致群体间的基因的交流。

迁移和突变的效应同样带来新基因进入群体, 或造成群体中基因频率的变化, 是一种定向的进化力量。

不同种族和不同民族的基因频率可有差异。迁移的结果使不同人群通婚, 彼此掺入外来基因, 导致基因流动, 可改变原来群体基因的频率。



如果迁入者和迁出者的比例各为 m ，设迁入者从其他群体随机地进入A群体，迁出者从A群体迁出至其他群体。且有 $0 \leq m \leq 1$ 。A群体中B、b基因频率分别为 p_0 、 q_0 ，迁入者的频率分别为 p_0' 、 q_0' ，迁移发生在群体中的个体进行生殖以前，迁入者在A中生殖，迁出者不在A内生殖。经过一代迁移后，混合群体中b基因频率为：

$$q_1 = (1 - m)q_0 + mq_0' = m(q_0' - q_0) + q_0$$
$$q_1 = q_0 - m(q_0 - q_0')$$

若每代 m 不变，连续 n 代之后的混合群体内的基因频率为：

$$q_n = q_0' - q_0'(1 - m)^n + (1 - m)^n q_0$$

可以证明：

$$(1 - m)^n = (q_n - q_0') / (q_0 - q_0') \quad (1)$$

式中： n ：迁移发生的世代数； m ：迁移速率或基因流动速率； q_n ：迁移 n 代混合群体中的b基因频率； q_0' ：迁入者的b基因频率； q_0 ：土著（或祖先）群体中的b基因频率



根据（1）式可计算 m ，当其他条件都已知时，由于迁移率比突变率要大很多，故迁移比突变更有效地改变群体的等位基因频率。

可以预计，迁移的最终效果将会是 $q_n = q_0'$ ，混合群体中的 q_n 与迁入者的 q_0' 相等，群体处于平衡之中。

这一过程也适合于下述情况，即当一个大群体分割成许多小的、部分隔离的群体，其间没有其他因素的干扰，那么亚群体间迁移的遗传学结果将会带来遗传混合（genetic admixture），使它们之间的基因频率都相同，等于整个大群体中的基因频率。

现以决定Rh血型基因座上的 R_0 为例说明基因流在人类群体中造成的后果。



设 R_0 基因频率在白人群体中的 $p=0.028$ ，而在黑人的祖先群体中 R_0 的基因频率 $p_0=0.630$ （或被视为黑人群体中的初始 R_0 基因频率）。已知，该国黑人的祖先来到这个国家大约为300年前（或以 $n=10$ 代表示），黑人群体中的 $p_n=0.446$ 。将上述数据代入公式（1）中，并将其中的 $q'_0 = p$ ， $q_n = p_n$ ， $q_0 = p_0$ 变换后可有：

$$(1 - m)^{10} = \frac{0.446 - 0.028}{0.630 - 0.028} = 0.694$$
$$1 - m = \sqrt[10]{0.694} = 0.964$$
$$m = 0.036$$

可见，白人群体进入黑人群体的基因流动速率相当于每代平均3.6%。经过10个世代的基因流，使这个国家中的黑人群体中的所有基因中剩下 $(1-m)^{10}=0.694$ 是来自黑人群体的祖先，而在黑人中已有 $1-0.694=0.306$ 或30.6%的基因是来自白人群体的祖先。



(2) 随机遗传漂变与群体大小

随机遗传漂变 (random genetic drift)：在有限个体的小群体中由于抽样的随机误差所造成的基因频率的随机波动。这是由S. Wright提出的关于群体遗传组成变化的一个重要理论，又称赖特效应 (Wright effect)。

遗传漂变是一个随机过程。一个新的突变基因 (或突变体) 很可能在群体中被随机丢失，或被固定，特别在那些小群体中此过程发生的速度很快。以标准差 (σ) 来测度生物学中的抽样误差与群体大小 (N) 的关系：

$$\sigma = \sqrt{\frac{pq}{2N}}$$



如果在 $p=q=0.5$ 的群体中，不难算出样本量为 $N=50$ 的标准差正好是 $N=5000$ 的10倍，充分说明样本愈小，抽样误差愈大。计算机模拟实验，证明了随机遗传漂变效应在不同群体中的结果（图19-1）。

从图中可见在每代保持 $N=25$ 的小群体中，经过漂变的随机变动，群体的等位基因频率在60代前已经被固定了（即 $p=1$ ）。而 $N=2500$ 的群体，经过150代的漂变作用仍能保持原来的 $p=0.5$ 频率，漂变的速度显然不及小群体快。随机遗传漂变在小群体中可能是主要的进化动力。

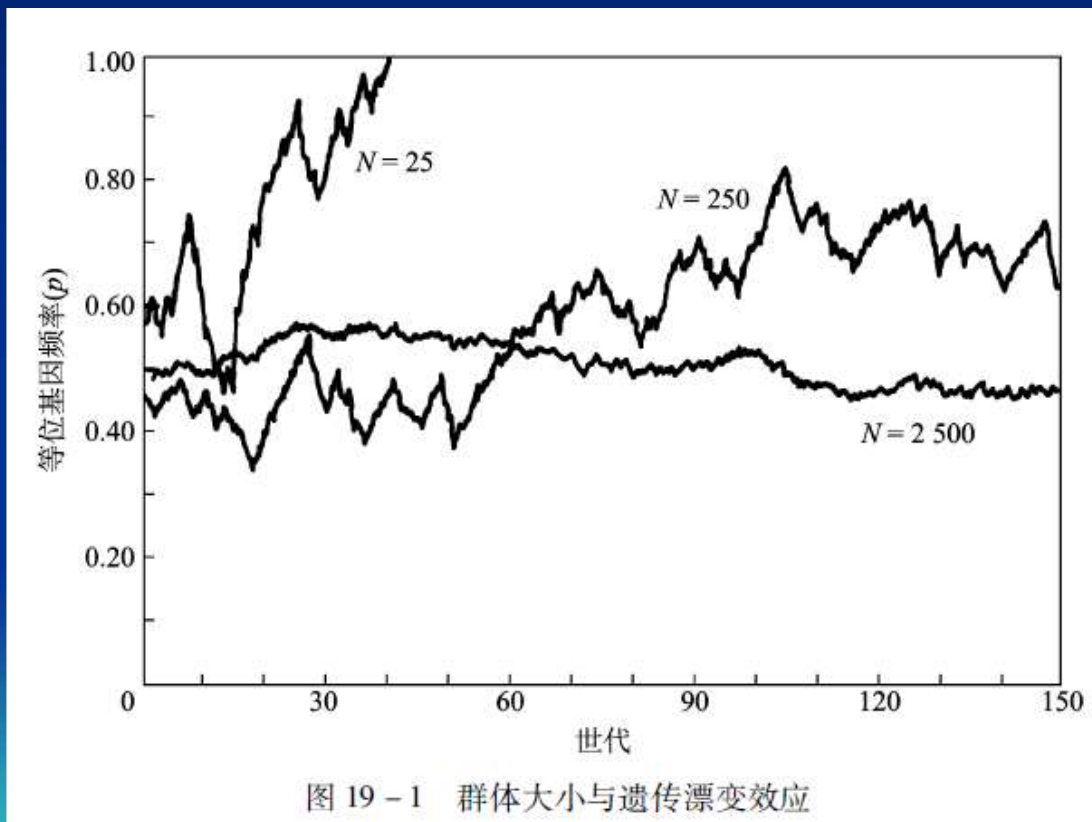


图 19 - 1 群体大小与遗传漂变效应



19.4 自然群体中的遗传变异及其检测

大多数生物的自然群体中具有大量的遗传变异，诸如表型变异、染色体结构变异、蛋白质及DNA序列多态性等，对其检测方法请详见e19-6。





19.5 物种及物种形成

19.5.1 物种及生殖隔离

(1) 物种的定义

经典的分类学对物种 (*species*) 的定义：物种是以表型特征为基础的，若两群生物的形态特征完全不同，则这两群生物就被认为属于不同的种。物种是生物分类学的基本单位。

生物学物种概念 (biological species concept, BSC) 1942年由Dobzhansky、Mayr和S. Wright等提出：“物种是一个具有共同基因库、能够自然地成功交配、并繁育出有生殖能力后代的一群个体所组成的一个群体，它们与其他群体是生殖隔离的”。

(2) 生殖隔离机制

生殖隔离机制 (reproductive isolation mechanism, RIM) 是生物防止杂交的生物学特征

分为两大类:

合子前隔离机制 (prezygotic isolation mechanism) 阻止不同群体成员间的相互交配, 阻止杂种合子的形成;

合子后隔离机制 (postzygotic isolation mechanism) 是一种降低杂种合子或杂种个体生活力或生殖力的生殖隔离 (表19-7)。

表 19-7 生殖隔离机制的分类

合子前隔离机制	<ul style="list-style-type: none"> ① 生态隔离: 群体占据同一个范围的地区, 但生活在不同的栖息地, 因此彼此不会相遇 ② 时间隔离: 交配期或盛花期在不同的季节或一天的不同时间 ③ 行为隔离: 雌雄间不存在性吸引或很弱 ④ 机械隔离: 交配或花粉传送受到外生殖器大小或形状不同或花的结构不同的阻挠 ⑤ 配子隔离: 雌雄配子不能互相吸引; 精子或花粉在动物生殖道或花的柱头上无生活力
合子后隔离机制	<ul style="list-style-type: none"> ① 杂种无生活力: 杂种合子无生活力或生活力降低 ② 杂种不育: 杂种不能产生有功能的配子 ③ 杂种衰败: F_2 或回交世代的杂种生殖力或生活力降低

合子前生殖隔离的例子：

金菊和翠菊在不同季节或一个季节的不同时间开花，卵子和花粉不是同时成熟，不能发生配子融合。



青蛙和蟾蜍，因不能同时达到性成熟而不能进行交互生殖。

合子后隔离的例子：

母马和公驴杂交所产生的骡子，是合子后隔离的经典例证。由于其杂交后代骡子不育，马和驴仍然是两个独立的物种。





阻止基因交流的隔离机制，还有**地理隔离**（geographical isolation）。分布在不同地区的群体，由于海洋、高山或距离遥远等地形地势的阻碍，使得彼此间没有交配的机会。地理隔离不依靠群体中任何遗传差异，遗传组成相同的群体可能在地理上被隔离（例如在孤岛上）。地理隔离往往与某种形式的生殖隔离密切联系在一起。



19.5.2 物种形成的方式

新的物种产生过程称为**物种形成 (speciation)**

在有性生殖的真核生物中，物种形成最为基本的方式是**渐进形式**：
由一个物种生发出多个新物种。

渐进的新种形成又可分为3种方式（表19-8）：

异域（地）物种形成（allopatric speciation）

邻域（地）物种形成（parapatric speciation）

同域（地）物种形成（sympatric speciation）

表 19-8 物种形成方式及其常见的遗传机制(引自 Brooker, 2001)

物种形成方式	遗传机制
异域(地):两个大群体被地理障碍隔开	在很长的时间内,众多微小的遗传差异慢慢累积逐渐导致了生殖隔离。一些遗传变异可能是适应性的,而另一些可能是中性的
异域(地):小群体的奠基者效应	遗传漂变和适应峰改变,可能导致新种的快速形成。如果群体被迁移到与以往生境差异很大的新环境中,自然选择将保留有利等位基因而淘汰有害的等位基因
邻域(地):两个群体占据重叠生境,杂交在部分情况下发生	由于杂交后代的适应性通常很差,等位基因的新组合或染色体重排可能迅速地限制了相邻群体间的基因交流
同域(地):在占据同一地域连续生境的群体内部,一个小群体演化为一个生殖隔离的新种	基因突变导致生殖隔离;例如突变可能影响配子识别。在植物中,四倍体的形成通常导致新种产生,因为四倍体与二倍体的杂交后代是三倍体,不可育



(1) 异域物种形成

当一个大群体中的一部分成员与其他成员在地域上隔离之后造成生殖隔离，形成新种。

例如，一道山脉的隆起隔开了一个处于低地域的群体；扩张中的冰川将一个群体割裂并阻碍了群体间的基因流。

例如：两个种的羚松鼠（antelope squirrel）分别在某大峡谷的两侧生存，南侧是哈里斯羚松鼠 *A. harrisis* (*Ammospermophilus harrisis*)，北侧是白尾松鼠 (*A. leucurus*)。它们是从一个共同祖先分化而来的，该祖先物种生活在大峡谷形成之前。随着时间的推移，在大峡谷形成之后，两个被隔离的群体逐渐独立地累积基因突变，通过自然选择和遗传漂变，最终导致两个表型上有差异的新物种形成（[图19-2](#)）。

而能轻易飞越大峡谷的鸟类并没有在峡谷两侧形成不同的物种。





(2) 邻域物种形成

物种的成员仅仅被部分地隔离，初始群体的地理分布区相邻接，地理隔离并不完全，群体间的个体在边界区有某种程度的基因交流，此类物种形成方式称为邻域物种形成。

例如，一座高山可能将一个物种分隔为两个群体，但在局部仍保留了一些通道，在这些仍然能使个体相互接触的地带中，两个群体的成员可以相互交配。诸如植物、陆地蜗牛、啮齿动物、蚱蜢、蜥蜴及很多不能飞行的昆虫都可能以这种方式形成新种。

在邻域物种形成之前，两个群体发生接触的通道被称为杂交区（*hybrid zones*）。在此区域内群体间的基因交流频率必须被限制在一定水平之下，新种才有可能发生。这可能由于边缘栖息地环境上的差别，使种群边缘的群体分化、独立，虽然没有出现地理隔离屏障，也能成为基因流动的障碍，在自然选择的作用下，逐渐形成生殖隔离机制而形成新种。





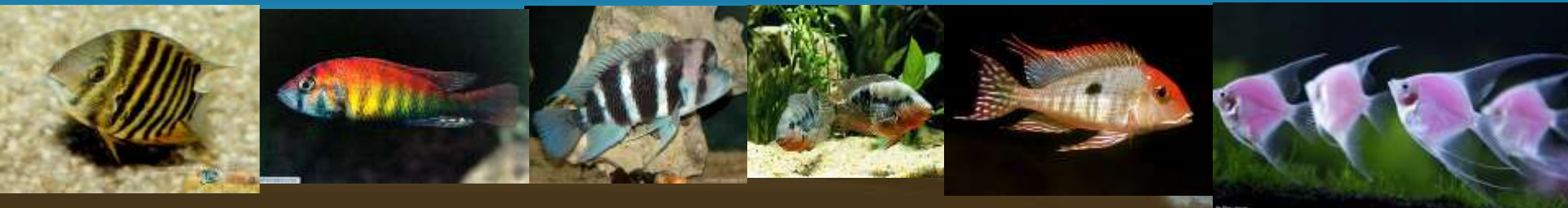
(3) 同域物种形成

同域物种形成 (sympatric speciation) 是指在同一地理位置从一个祖先物种形成后代物种的现象。

例如：在东非马拉维湖、坦噶尼喀湖和维多利亚湖三大湖泊里生活着的热带淡水鱼——**丽鱼科 (Cichlidae)** 中的各种丽鱼 (cichlid) 是从**一个单独的原始物种**，在一个有限的地域内经过暴发式的适应辐射而形成的具有高度多样性的、大量的新物种。

现今，马拉维湖中生存着800到1000种、坦噶尼喀湖内大约有250多种、维多利亚湖中具有500多种丽鱼。

在同一水域中的数百种丽鱼的快速形成的关键因素之一是性选择——不同体表颜色的雄鱼选择具有特殊颜色的雌鱼交配，以致形成了在生殖上相对隔离的、不同群体对不同食物的适应，最终导致新物种的形成





植物多倍体的形成导致突然地发展出生殖隔离，跳跃式地形成新物种。例如，一个荨麻新种 (*Galeopsis tetrahit*) 是来源于两个二倍体物种 *Galeopsis pubescens* 及 *Galeopsis speciosa* 的异源四倍体。这两种二倍体植物各有 16 条染色体 ($2n=16$)，而新种 *G.tetrahit* 含有 32 条染色体。此异源四倍体之间的杂交后代是可育的，因为所有的染色体在减数分裂中都能形成同源配对并正确地分离，其子代仍为四倍体。如将此异源四倍体与二倍体杂交，后代将形成不育三倍体，因为它必然会产生有缺陷的配子。这样的杂种不育性将异源四倍体与二倍体在生殖上隔离开来。异源八倍体小黑麦新物种的形成是其中另一重要例证，详见第 12 章。



达尔文生物进化学说：认为相似的生物都起源于一个共同的祖先，都可以发生变异，每个个体都会携带着不同的变异，**生物进化的动力源于“自然选择”**，那些具有有利变异的个体更有可能生存下去，把基因传递给后代；而那些产生有害变异的个体，可能会在自然选择中被淘汰。生物的进化是一个逐渐的、连续的过程。**该学说指出生物进化的主导力量是自然选择，生物进化是遗传、变异和选择三者综合作用的结果。**

综合进化学说或新达尔文学说：该学说认为，**突变是遗传变异的根本原因**，但自然选择却在决定群体的遗传组成，并在基因替换过程中起着**决定性的作用**。选择被视为是驱动进化过程的唯一力量，其他因素如突变、随机遗传漂变被认为最多只有次要作用。**突变、选择和隔离是物种形成和生物进化的机制。**





分子进化 (molecular evolution)：是指DNA、RNA和蛋白质水平上的进化过程。其研究目的是通过比较、分析DNA、RNA和蛋白质分子序列、结构与功能的改变来探索群体进化速率、物种间的亲缘关系、进化的过程和机制。

分子钟 (molecular clock)：在生物的各个系谱中，明显的序列改变的恒定速率。

蛋白质序列实验数据显示功能相同的血红蛋白和细胞色素c在不同物种间有若干氨基酸的差异，差异的数目随着它们从共同祖先分化的时间而递增（[图19-3](#)）。

✓ 分子进化的中性学说 (neutral theory of molecular evolution)

1968年由日本学者木村资生 (M. Kimura) 等提出, 该学说认为: **分子水平上的大多数突变 (DNA突变、氨基酸的替代, 以及物种中随时间进行的大多数变异) 不是由达尔文的自然选择, 而是由选择上呈中性或近中性的突变等位基因的随机遗传漂变被固定下来的。**

中性学说的本质并不是强调分子突变是严格意义上的选择中性, 而在于它们被固定的命运在很大程度上是由随机遗传漂变决定的。换言之, **在分子进化过程中, 选择压是如此微弱, 以致突变压 (mutation pressure) 和随机遗传漂变起着主导的作用。** 该理论的提出, 是进化学说概念上的一次重大飞跃。



DNA分子随时间发生的中性替代由遗传漂变被固定下来的速率可以按下列方法计算：

设：中性替代速率= k ；每一世代中，每个基因拷贝在一个位点上发生新突变的突变率= μ ；一个中性等位基因最后被固定的概率= x 。

在有 N 个二倍体组成的群体中，每个基因座（常染色体上）有 $2N$ 个等位基因， $k=2N\mu x$ ；若等位基因是中性的，所有基因被固定下来的概率都相等，即 $x=1/(2N)$ ；

$$\text{故 } k=2N\mu \cdot 1/(2N) = \mu$$

由此可知，在群体内的每一世代中，遗传漂变作用将会固定下 μ 个中性突变，即中性等位基因的替换率直接等于它们的突变率，它与群体大小及其他任何参数均无关。如果突变率保持恒定，则分子进化速率也将保持恒定，反之，中性等位基因替代速率的估计直接提供了中性突变率的估计。



19.7 新基因和蛋白质功能的起源

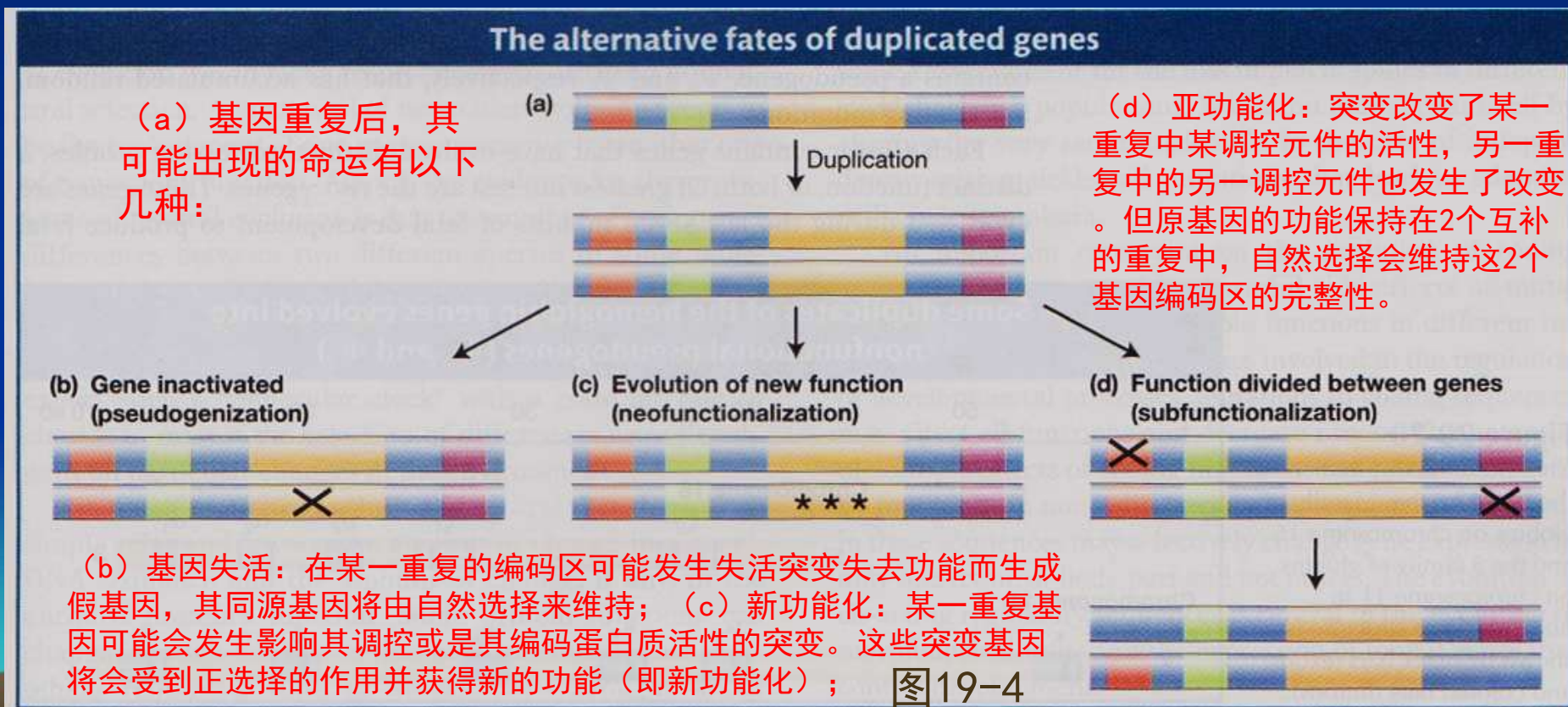
(1) 基因数目的扩增

基因数目扩增方式有如下几种：

- ① **多倍体化** 形成多倍体，在植物物种进化过程的作用十分重要。
- ② **基因重复** 基因组中某个特定基因复制增加一份，可产生两个相同的基因。
- ③ **转座** 当转座因子转座时，携带原位点处额外的片段，并在基因组中新的位点插入这段拷贝
- ④ **反转录转座** 在许多动物的基因组中包含这种转座因子。在很偶然的情况下，宿主基因组的mRNA反转录为cDNA并插回基因组中，导致缺乏内含子的基因重复。

(2) 重复基因的命运

基因组中重复的基因实际上是备用的遗传元件，它们都可以独立进化出各种新的功能（称为**新功能化，neofunctionalization**）。简便起见，只考虑基因组中一个发生在某基因整个编码及调控区上的重复（图19-4），来了解新基因的命运及其新功能的进化方式。





19.8 人类进化概述

19.8.1 人类的近亲及其分子生物学证据

人类与黑猩猩（*Pan troglodytes*）有共同祖先。但是，人类进化史上最重要且最难确定的问题是黑猩猩与人类究竟是在什么时候分道扬镳的？遗传学家研究人类及黑猩猩的DNA分子证据，最早按分子钟（molecular clock）推算出人类祖先是在距今400万年至600万年前与黑猩猩在进化树上分歧的。

Figure 20.5 Humans diverged from an ancestor shared with chimpanzees about 6 million years ago. Representatives of primates alive today: (a) orangutan, (b) gorilla, (c) chimpanzee, and (d) human.



(a)



(b)

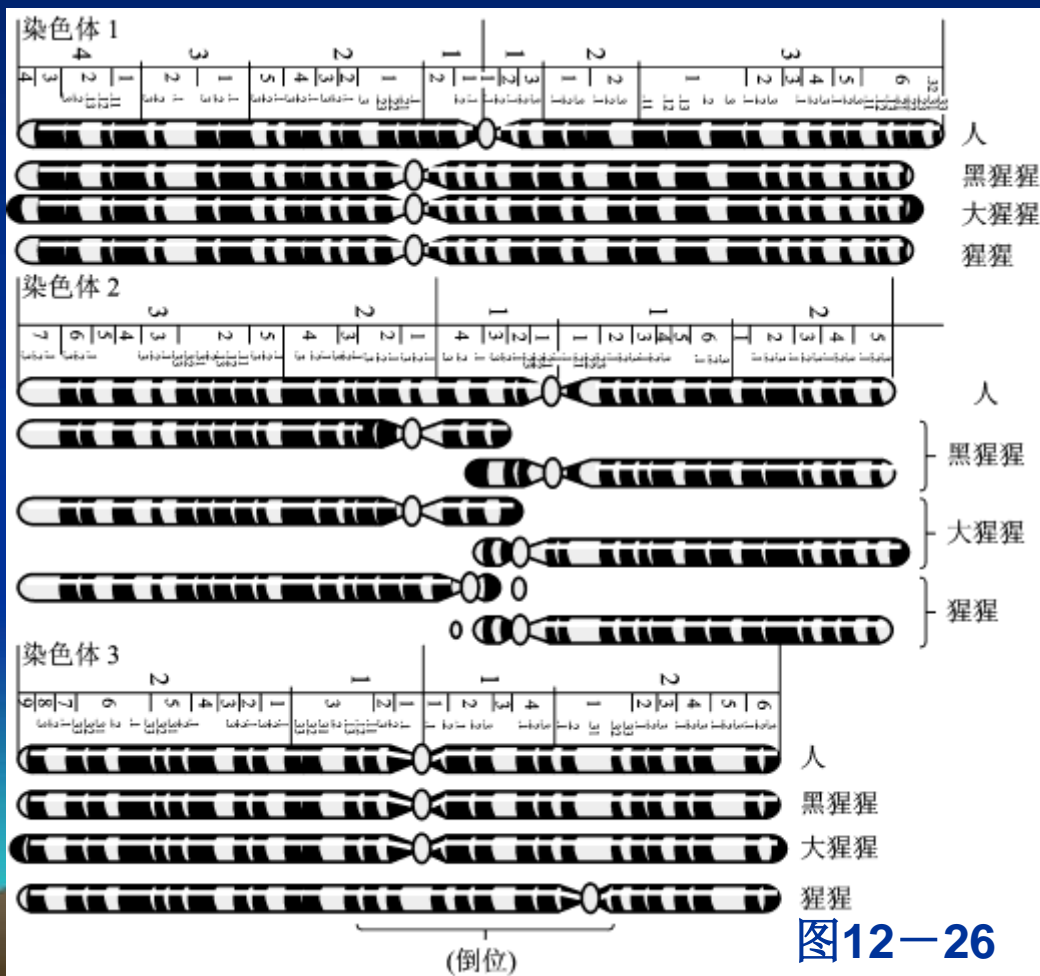


(c)



(d)

细胞遗传学研究表明，从一个共同祖先到人科的进化中，人类系谱中至少发生了一次罗伯逊易位（着丝粒融合）。此外，图12-26中显示了人类的第3染色体与黑猩猩、大猩猩的带型很相似，但在猩猩中则显然发生了一个大片段的臂间倒位，因为在其长臂靠近着丝粒区出现了带纹的重排。



黑猩猩



大猩猩



猩猩

图12-26



近几年，基于新的分子钟的分析，对于人类历史的真正开端获得了令人兴奋的新证据：

Kong扫描78个儿童及其父母的基因组，发现平均每个孩子带有36个突变，比之前估计的平均72个减少了一半，即分子钟比之前估算的要慢很多。

Langer-graber对8个野生黑猩猩群中226只新生黑猩猩的分析认为：黑猩猩的平均生育年龄为24.5岁（即黑猩猩完成一个世代所需的时间）。根据新的分子钟估计出人类与黑猩猩在进化树上分歧的时间至少在700万年以前，甚至可能推前更远至1300万年！



之后，Lovejoy认为人-猿分歧时间可以远至1300万年之前，他指出在1000万~1500万年前的地球上分布着各种已经具有若干现代人类才有的解剖学特征的猿类。

Lovejoy分析了85000个冰岛人的、记录突变率更可靠的微卫星DNA的突变速率，推断人与黑猩猩分歧时间更为保守的数字仅为750万年以前！这意味着人类起源的历史比人们曾经所知的要更加“古老”！换言之，我们人类的近亲，离我们更加“遥远”！因此，更多的类人物种将进入到人类进化的大谱系中来（图19-5）。

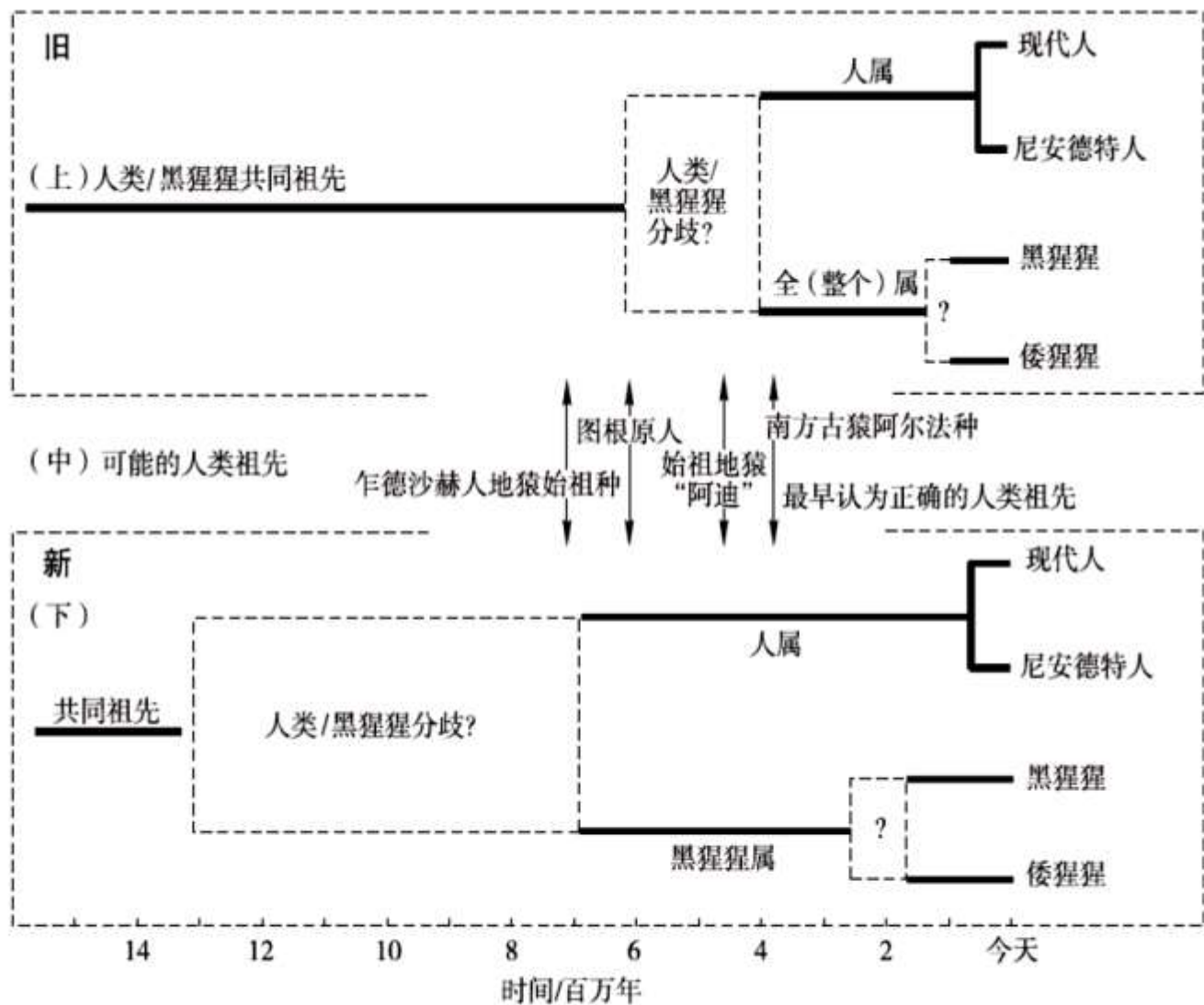


图 19-5 人类进化路线图(仿自 Brahic, 2012)

(上)400 万~600 万年前人类祖先与黑猩猩的分歧 (中)人类可能的祖先 (下)最新发现人-黑猩猩分歧时间推前至 700 万~1 300 万年前



19.8.2 现代人的进化之路

基于新的分子钟的分析，确认并更新了人类起源的谱系，也将起源于非洲的现代人走出非洲的时间推移到距今9万年~13万年。

我们现代人是怎样从祖先进化而来？

进化过程中曾发生过怎样的事件？

哪些事件对于现代人的成功进化起了怎样的作用？

这是科学家需要回答的一些重要理论问题。为此先后有人提出以下观点：

第一种观点是：“走出非洲模型”（out of africa model）

20世纪末，古人类学家依据化石证据认为解剖学意义上的人属早期成员起源于撒哈拉以南非洲单一地区的直立人。随后，从非洲迁徙到旧大陆的其他各地区，从而取代了当时尚存于当地的古人类物种。也称为“非洲取代模型”（african replacement model）



第二种观点认为：现代人的祖先到达新地区后，和其他古人类群体存在偶发性的基因交流——“**杂交模型**”（hybridization model）。

第三种模型为“**多地区连续进化模型**”（multiregional evolution model），认为各个古人类群体转变为现代人是逐渐发生的。生活在非洲、欧洲或大洋洲的人在迁移过程中相互间发生了基因的交流，使得有利于各人群适应当地环境的特征得以保留了下来。

在上述第三种模型的基础上演生的观点认为：现代人的特征主要来自非洲祖先的同化过程，这便是第四种被称为所谓的“**同化模型**”（assimilation model）。





近代，在分子进化研究方面，由于人类线粒体基因组（mtDNA）进化速率快、基因组小、并是母系遗传，利用其研究现代人的进化最为深入。

例如：Wilson等从200多个代表许多不同种族（或民族）的个体中提取的mtDNA，通过序列分析与构建进化树发现：**所有现代人群体的mtDNA都来源于20万年前生活在非洲的一位女性个体（被昵称为人类共同的线粒体母亲）。**

此后，德国科学家Svante Pääbo于1997年，研究发现尼安德特人与现代人的mtDNA存在显著不同。依mtDNA分子谱系分析可以肯定现代人取代了尼安德特人。



科学家对现代人和其他古人类的DNA序列比对的研究表明，来自非洲的现代人在他们地走出非洲的过程中，在不同时期和地区与其他古人类相遇时发生了广泛的基因交流。

早期的现代人与其他古人类物种间的基因交流所带来的“混血”（图19-6）无疑对现代人的成功进化具有至关重要的作用





研究发现，现代人许多新特征是在近8000~10000年迅速出现的，这说明人类一直在进化着！例如：

(1) **乳糖酶** 编码乳糖酶的*LCT*基因与乳糖耐受性相关。研究表明，该基因在最近的进化历程中至少发生过5次突变。第5次突变最常见，从爱尔兰到印度的广大地区的成年人体内的*LCT*基因一直保持活性。该新突变受到正选择的作用而以指数式快速扩散，其中在北欧人群中频率最高。突变后的乳糖酶基因在欧洲人群中独立进化，而亚洲和拉丁美洲的成年人大多数携带着原始的乳糖酶基因，以致使不能消化乳糖。

(2) **叮叮** 研究发现，“干型”叮叮是*ABCC11*基因在近代发生的一个隐性突变所致，此基因突变是2万~3万年前才出现的。*ABCC11*基因还影响人体的汗腺——使得腋窝经常潮湿且带异味（俗称腋臭症），“干型”叮叮者体味正常！



(3) **唾液淀粉酶** *AMY1*基因编码唾液淀粉酶 (salivary amylase)。人类基因组中*AMY1*基因的拷贝数尤其多，使得依靠淀粉含量高的谷物类为主要食物来源的人们能够消化更多的淀粉。

(4) 其他新特征的进化

以下诸多新特征都是最近才以惊人的速度进化而来的：东亚人常见的直而浓密的黑头发是在过去3万年间才开始出现的，是由*EDAR*基因突变产生的新特征；人类的祖先的皮肤、毛发、虹膜的颜色先前都是深色的，发生基因新突变后，使得在一定程度上都变浅了，诸如在所罗门群岛人群的黑发变成金黄色是由于*TYRP1*基因突变所致；*MC1R*基因突变带来了“红发”人（全球不到1%），他（她）们大多数分布在英国、爱尔兰、澳大利亚和荷兰等地；90%的欧洲人的基因组中的*SLC24A5*突变基因给当地人带来了浅色的皮肤；9000年前*HERC2*基因的突变出现了蓝色的虹膜（蓝眼睛）。



未来人类将进化成什么样子？过去数千年来，不同地区的人群各自经历着不同的进化过程，但仍然保留着惊人的共性，新的适应性变异无法完全取代原有的基因，从祖先基因组中继承下来的大多数“旧基因”仍然保留着。但是所有的地球人共享着一个巨大的基因库，每年、每月甚至每时每刻都有数以百万计的人群在不同国家或地区迁徙着，遗传信息正以前所未有的速度进行着广泛的交流与混合，毫无疑问，可以预计未来的人类将会出现更多的新特征。



武汉大学

Wuhan University

谢谢！

作业： p436-1、 p437-5

本章结束





关于遗传学复习考试

1、全面复习：基本概念、基本作图方法、

基本运算公式、基本解题方法

2、要求掌握：

1)、概率

2)、二项式展开

3)、 $AaBa \times AaBa$ F_2 表型比例、基因型比例

4)、多倍体配子（三倍体，四倍体）的产生及其比例

5)、重组频率 = $\frac{\text{重组组合}}{\text{亲组合} + \text{重组组合}} \times 100\%$

重组作图

6) 四分体分析: 判断依据 PD>>NPD 连锁

$$R_{a-b} = (1/2T + NPD) / (PD + NPD + T) \times 100\%$$

第二次分裂分离子囊数 (或交换型子囊数)

$$\text{基因与着丝粒的重组率} = \frac{\text{第二次分裂分离子囊数 (或交换型子囊数)}}{\text{子囊总数}} \times 1/2 \times 100$$

$$\text{或 } F(\text{着丝粒-基因}) = \frac{M_{II} \times 1/2}{M_I + M_{II}} \times 100\%$$

7) 共转化频率的计算:

$$\frac{\text{Number of } a^+b^+ \text{ cotransformants}}{\text{total number of } a^+ \text{ transformants}} \times 100\%$$

8) 细菌: 内基因子、外基因子, 相反的重组子不成在

$$R_f = \frac{\text{重组型 (+-) 和 (-+)}}{\text{亲本型 (++) + 重组型}}$$

9)、共转化频率与图距的相互关系的数学表达式:

$$X=(1-d/L)^3$$

10)、Hardy-Weinberg平衡:

$$(p+q)^2 = p^2 + 2pq + q^2$$

以及基因频率和基因型频率的估算

11)、近交系数的估算

12)、孟德尔遗传定理、连锁互换定理的运用

13)、 $V_p = V_A + V_D + V_E + V_I$

广义遗传率(力)和狭义遗传率

14)、核外遗传与核内遗传的特点、关系

15)、遗传作图: 请总结各种作图方法



3、题型：

1)、名词解释 **30**分 (10个名词, 每个3分)

2)、简答题 **30**分 (5个简答题, 每个6分)

3)、计算、分析作图题 **40**分 (4个题：
每题各**10**分)

题目简单、覆盖面广、综合性强